HSR Health Services Research

DOI: 10.1111/1475-6773.12449
METHODS ARTICLE

Health Insurance Dynamics: Methodological Considerations and a Comparison of Estimates from Two Surveys

John A. Graves and Pranita Mishra

Objective. To highlight key methodological issues in studying insurance dynamics and to compare estimates across two commonly used surveys.

Data Sources/Study Setting. Nonelderly uninsured adults and children sampled between 2001 and 2011 in the Medical Expenditure Panel Survey and the Survey of Income and Program Participation.

Study Design. We utilized nonparametric Kaplan–Meier methods to estimate quantiles (25th, 50th, and 75th percentiles) in the distribution of uninsured spells. We compared estimates obtained across surveys and across different methodological approaches to address issues like attrition, seam bias, censoring and truncation, and survey weighting method.

Data Collection/Extraction Methods. All data were drawn from publicly available household surveys.

Principal Findings. Estimated uninsured spell durations in the MEPS were longer than those observed in the SIPP. There were few changes in spell durations between 2001 and 2011, with median durations of 14 months among adults and 5–7 months among children in the MEPS, and 8 months (adults) and 4 months (children) in the SIPP.

Conclusions. The use of panel survey data to study insurance dynamics presents a unique set of methodological challenges. Researchers should consider key analytic and survey design trade-offs when choosing which survey can best suit their research goals.

Key Words. Health insurance, survival methods, duration models, survey methods

Insurance coverage estimates from household surveys are an important input into the policy making process. For example, state estimates on the share of the nonelderly population without health insurance have been used by Congress to set funding levels for the State Children's Health Insurance Program, and they also serve as a key input to formulas that determine uncompensated care payments to hospitals (Davern and Blewett 2007; Graves 2012). More recently there has been interest in using survey data to assess how the Affordable Care Act (ACA) is changing the U.S. health insurance landscape (Carman and Eibner 2014; Long, Zuckerman, and Kenney 2014; Sommers et al. 2014).

The vast majority of research on insurance utilizes estimates on the share of the population with or without coverage at a point-in-time. However, for policy making and evaluation purposes it is also important to understand the dynamic processes generating coverage patterns. For example, successful ongoing implementation of the ACA will depend on identifying and enrolling people who remain without insurance. If those remaining uninsured are without coverage for short periods, then greater emphasis could be placed on outreach efforts that highlight the use of health insurance exchanges for transitional coverage. By contrast, if those remaining uninsured are predominantly in long spells, then achieving further coverage gains may require reevaluating the generosity of subsidies to make insurance more affordable (Graves and Swartz 2012, 2013).

Obtaining accurate estimates of insurance dynamics is challenging due to methodological issues researchers face when using longitudinal household surveys. These issues—including attrition, censoring, truncation, seam bias, and the choice of survey weights—are in addition to the well-documented challenges of estimating coverage rates at a point-in-time (Swartz 1986; Davern et al. 2004, 2007; Call, Davern, and Blewett 2007; Call et al. 2008). To date, however, an overview of these methodological issues and how they manifest in commonly used panel surveys does not exist in the health services literature.

Our study's primary objective is to provide researchers with an overview of the methodological issues faced when using longitudinal survey data to study insurance dynamics. We use this discussion as an opportunity to highlight and test how both survey design and analytic choices may affect estimates of uninsured spell durations across the two most widely used longitudinal surveys, the Survey of Income and Program Participation (SIPP) and the Medical Expenditure Panel Survey (MEPS). Specifically, using 2001–2011 SIPP and MEPS data, we show how estimates of uninsured spell durations differ across the two surveys and compare the robustness of estimates to different methodological approaches.

Address correspondence to John A. Graves, Ph.D., Department of Health Policy, Vanderbilt University School of Medicine, 2525 West End Avenue, Suite 1200, Nashville, TN 37203-1738; e-mail: john.graves@vanderbilt.edu. Pranita Mishra, M.P.P., is with the Department of Health Policy, Vanderbilt University School of Medicine, Nashville, TN.

Background

Most studies of insurance dynamics draw upon a class of methods known as survival or event history methods. In contrast to approaches that focus on the population distribution of people with or without insurance, event history methods focus on the distribution of insured or uninsured spells themselves. That is, the research emphasis is on understanding (un)insurance as an event rather than as a characteristic.

Quantities of research interest in a study on insurance dynamics could be quantiles in the distribution of spell durations (e.g., the median uninsured spell length), estimates of how the likelihood of leaving an uninsured spell evolves as a function of time, or the influence of key variables (such as exposure to a policy change) on the probability (i.e., "hazard") of losing or gaining insurance.

The logistical and budgetary barriers to recruiting a nationally representative cohort are substantial, so estimating these quantities often requires retrospective use of household surveys. Previous research has demonstrated that cross-sectional insurance estimates differ across commonly used surveys (Call, Davern, and Blewett 2007; State Health Access Data Assistance Center, 2010). There are a number of explanations for this variation, including differences in survey reference periods, undercounts of individuals enrolled in Medicaid, and alternative techniques for data editing and imputation (Davern et al. 2004, 2007; Call, Davern, and Blewett 2007; State Health Access Data Assistance Center 2010).

The use of panel surveys to study coverage dynamics presents a unique set of additional challenges. As detailed below, these challenges range from the influence of sample attrition to the theoretical properties of spells and how they are sampled in a population. While beyond the scope of our study, researchers also face methodological choices in whether to assume a parametric form for the underlying distribution of spells, and in whether and how to account for unobserved heterogeneity in their models (Lancaster 1992).

Past research efforts differ in how they address these challenges. Swartz and McBride (1990) utilize nonparametric life-table methods on incident uninsured spells (i.e., spells with observed beginnings) in the 1984 SIPP and find median durations of about 4 months. Later parametric modeling using the 1984 SIPP was influential in highlighting the importance of length bias (discussed below) when the analytic sample includes uninsured spells already in progress at the beginning of the survey (Swartz, McBride, and Marcotte 1993). Similarly, research by Cutler and Gelber (2009) utilizes competing risks models on SIPP data to examine how rates of exit from uninsured spells to private and public coverage evolved from 1983–1987 to 2001–2004. Most recently, Graves and Swartz (2012) utilize nonparametric survival methods and SIPP data to demonstrate how the 2006 Massachusetts reforms impacted the duration of uninsured spells. Our study adds to this small but influential literature by highlighting recent trends in uninsured spell dynamics and by demonstrating how survey features and analytic choices may impact estimates in two of the most commonly used surveys.

Data and Methods

Data

We utilize self-reported data on insurance coverage in each month from the 2001, 2004, and 2008 panels of the SIPP and panels 6 through 15 of the MEPS. The SIPP is a nationally representative panel survey of U.S. households conducted by the Census Bureau (U.S. Census Bureau, 2001). Responding households are interviewed in person at baseline and then by phone or in person every 4 months for up to 4 years.

The MEPS is a household survey sponsored by the Agency for Healthcare Research and Quality. It utilizes an overlapping panel design with a new 2-year panel drawn annually from the sampling frame of the previous year's National Health Interview Survey (NHIS) (Cohen 1997). This design stands in contrast to the SIPP, which has a single panel in the field at any given time. MEPS households are surveyed in person five times, with interviews spaced, on average, about 6 months apart.

Our analytic sample includes nonelderly adults ages 18–61 years who do not age into Medicare while in the survey, and children under 18 years old. Because the 2008 SIPP was not fielded until September 2008, we exclude MEPS respondents whose uninsured spells began between January and August 2008. Similarly, MEPS data beyond 2011 were not available at the time of our study, so our sample excludes SIPP spells that began after December 2011.

To facilitate comparisons across surveys we assign observations in each MEPS panel to one of three periods depending on when their uninsured spell began: 2001–2003, 2004–2007, and 2008–2011.

Methodological Considerations

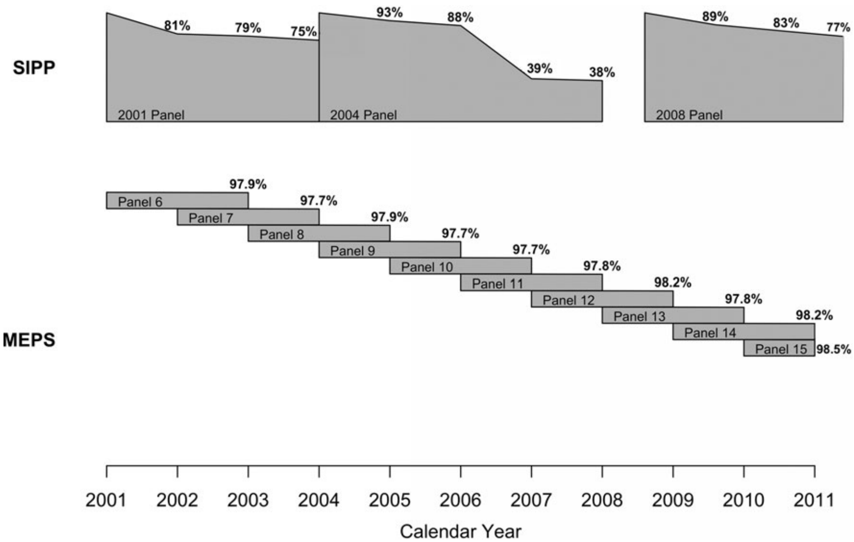
Unit Nonresponse and Attrition. Nearly every longitudinal survey suffers from sample attrition, and the SIPP and MEPS are no exception. SIPP response rates are about 90 percent in each round, yielding an overall loss of 25 percent in the 2001 panel, 62 percent in the 2004 sample, and 23 percent in the 2008 panel (U.S. Census Bureau, 2009).[1] By contrast, there is much less attrition in the MEPS, with 2-year cumulative attrition rates averaging between 1.5 and 2.3 percent in our nonelderly sample. This difference between surveys is explained in part by the longer panel length of the SIPP (3–4 years) versus the MEPS (2 years). However, even the 2-year SIPP nonresponse rate (roughly 13 percent in our sample) is much higher than the MEPS.[2] This difference likely stems from the fact that the MEPS is sampled from respondents to the previous year's National Health Interview Survey (NHIS). That is, because MEPS respondents have previously participated in a household survey, they may have observable (and unobservable) characteristics that make them significantly less likely to drop out.

To illustrate these issues and how they may affect estimates of insurance dynamics, Figure 1 plots each panel sample size over time. Each polygon represents a distinct survey panel, with the height scaled proportionately to reflect the contribution of that panel to our overall sample. We adjust the height of each polygon downward in each year based on attrition.

Several important observations emerge from Figure 1. First, the nonoverlapping nature of the SIPP indicates that estimates of insurance transitions are estimated using the sample of households remaining after each round. By contrast, the first year of each new MEPS panel overlaps with the second of a previous MEPS panel. This overlapping design results in a new MEPS panel sampled in each calendar year.

These contrasting sampling designs could impact estimates of uninsured spell durations. Previous research on SIPP respondents matched with administrative income data has shown that young (age: 18–24) individuals with nonpositive earnings are more likely to drop out (Czajka, Mabli, and Cody 2008). Moreover, both the MEPS and SIPP oversample individuals likely to participate in public programs. These individuals have been shown to have less stable housing situations relative to the general population, and therefore may be more likely to drop out of the survey (Kauff, Olsen, and Fraker 2002). As income, age, and participation in public assistance programs are well-known correlates of insurance coverage (Fronstin 2010), samples that restrict to individuals observed for the full

Figure 1: This Figure Plots the Relative Contribution of Each Survey Panel for the 2001–2011 Study Period



Notes. In the figure, each polygon represents a distinct survey panel, with the initial height scaled to reflect the original panel sample size. As each panel progresses over time (measured along the *X*-axis), the height of each panel's polygon is adjusted downward to reflect attrition from the original sample. The gap between the second and third SIPP polygons represents the 7-month lag between the end of SIPP 2004 and the start of SIPP 2008.

panel may yield biased estimates if insurance transitions are drawn from a sample of households with systematically different traits.[3] For example, Short, Graefe, and Schoen (2003) find evidence that the sample of fully observed 1996 SIPP respondents had a lower point-in-time uninsured rate than the full sample. This finding is consistent with the hypothesis that respondents with less stable income and housing situations are more likely to drop out. We discuss the trade-offs to restricting samples to fully observed respondents in the section on survey weights below.

Censoring and Truncation. When studying insurance dynamics it is important to consider that spell durations observed in longitudinal surveys are not random samples from the distribution of spells. This is true even if the underlying data are drawn from a random sample of households. For instance, if an individual's longitudinal data end before their spell completes, that spell is right-censored. To obtain unbiased estimates of spell durations, researchers must take care to account for right censoring in their analysis (Swartz and McBride 1990; Swartz, McBride, and Marcotte 1993; Klein and Moeschberger 2003). Otherwise, treating the last observation month as the "end" of a spell will lead to underestimates of spell durations.

Fortunately, most event history methods are designed to account for right censoring. Moreover, the presence of censoring will not lead to biased estimates as long as the censoring mechanism is unrelated to the event of interest (Cox and Oakes 1984). The presence of a large amount of right censoring can still pose a hurdle, however, as it can potentially restrict the set of quantities that can be estimated. For instance, obtaining nonparametric estimates of median spell durations may not be possible if well over half of the spells have no observed ending. This is a particularly important consideration for researchers considering the use of surveys (like the MEPS) with shorter follow-up periods.

In contrast to approaches that include right-censored spells, the inclusion of spells already in progress at the beginning of a survey poses more significant methodological trade-offs. As detailed by Swartz, McBride, and Marcotte (1993), and as illustrated in other contexts (Wolfson et al. 2001; Asgharian, M'Lan, and Wolfson 2002; de Una-Alvarez 2004), these spells known to be length-biased. Unlike left censoring, which would occur if an uninsured spell concludes before observation of a given individual begins, spells in progress at the beginning of a survey are left-truncated as the spell is observed from a delayed entry time (Boudreau 2003).[4] By definition, these spells have not concluded by the time the individual enters the survey; thus, they are disproportionately representative of individuals in longer spells. Failure to account for this statistical feature may lead to biased estimates of hazards and spell durations (Wolfson et al. 2001).

When considering whether to include left-truncated spells, researchers face trade-offs between efficiency gains from including more spell observations and the additional modeling assumptions required to obtain unbiased estimates. Including left-truncated spells in the 2001–2011 MEPS, for example, increases the sample size by about 48 percent, whereas in the SIPP the sample increases by 39 percent. While including these observations can improve estimation by adding valuable information for individuals in longer spells, researchers must consider the additional modeling assumptions needed to account for the theoretical properties of length-biased spells (Boudreau 2003). When using nonparametric methods, for example, the inclusion of left-truncated spells requires stationary assumptions on the underlying data-generating processes if no information is available for when the spells began (de Una-Alvarez 2002, 2004). That is, researchers must assume that the process generating these spells is in equilibrium (Boudreau 2003). In the context of studying spells without health insurance this may be unreasonable as the factors that influence the incidence and duration of periods with and without coverage can fluctuate for any number of reasons, including changes in public policies and in response to macroeconomic conditions.

One potential solution is to adopt a parametric approach and use information from freshly observed spells to model the data-generating process. In practice, this amounts to adopting a maximum likelihood-based approach with a separate likelihood function defined for left-truncated spells (Lancaster 1992). Even under that approach, however, equilibrium assumptions are necessary when there is limited or no information on when each spell began (Boudreau 2003).

The alternative to a fully parametric approach is to avoid left-truncated spells altogether by simply restricting the sample to incident spells—that is, spells with observed beginnings. We adopt this approach here, as our objective was to show how estimates from the SIPP and MEPS differ while making the minimum necessary assumptions on the underlying data-generating processes.

Survey Weights. While the use of survey weights is a disputed topic in the survey sampling literature (Lohr 1999), the use of sampling weights for the analysis of insurance coverage is advised as most surveys are not simple random samples of U.S. households. More generally, the use of weights is recommended whenever the sampling design is "nonignorable" with regard to the event or outcome of interest (Lohr 1999). As both the SIPP and MEPS oversample low-income and minority populations, and as both income and race/ethnicity are widely regarded as correlates of insurance, the use of sampling weights is advised. The choice researchers face is therefore not whether to use weights, but which set of weights to use.

The public release of panel surveys like the SIPP and MEPS includes a variety of weights, including cross-sectional (i.e., point-in-time) and longitudinal weights. Cross-sectional weights allow for nationally representative analyses of survey respondents at a point-in-time, while longitudinal weights facilitate representative analyses of a population over a defined time period (e.g., a specific calendar year or over a multiyear panel) while also correcting for attrition from the survey.

Both point-in-time and longitudinal weights include a "base" component to account for the differential probability of initial selection. Base weights are typically lower for respondents in groups sampled at higher rates. In addition, most survey weights also incorporate further adjustments for unit nonresponse and for departures of survey estimates from known population totals (Lohr 1999).

The public release of the SIPP includes both calendar month cross-sectional weights, as well as longitudinal weights defined for specific calendar years and for the entire multiyear panel. The MEPS, by comparison, includes a full-year weight that allows for point-in-time analyses of the population at any time within a calendar year (and, for some variables, on December 31). The MEPS also includes a longitudinal weight for representative analyses of the population over the 2-year panel.

While the availability of longitudinal weights may appear appealing for studies of insurance dynamics, these weights can be avoided if cross-sectional weights are available (U.S. Census Bureau 2008a,b). While counterintuitive, this may be a preferred strategy for two reasons. First, the added value of longitudinal weights (i.e., adjustments for nonresponse over time) may be redundant as most event history models already account for random right censoring. Second, longitudinal weights are defined only for respondents fully observed over the longitudinal time horizon. Thus, utilizing longitudinal weights effectively drops all other observations from the estimation sample, reducing effective sample sizes. By definition, to maintain estimates that are nationally representative, this increases the weight placed on fully observed respondents—and any systematic differences between these respondents and those who drop out must be accounted for either in the construction of the longitudinal weights themselves, or through the use of controls in a regression model.

As noted by the Census Bureau, an alternative approach is to weight observations using the cross-sectional weight corresponding to the month when an individual's spell begins (U.S. Census Bureau 2008a,b). For the SIPP, this amounts to utilizing the calendar month weight; for the MEPS, it amounts to using the full-year weight for the panel year in which the spell begins. A key advantage of this approach is that researchers can maximize sample sizes while also correcting for the complex sampling design. In the 2004 SIPP, for example, this method yields a final sample size that is four times larger than using the longitudinally weighted full panel sample. We investigate the sensitivity of the use of longitudinal versus cross-sectional weights for uninsured spell estimates below.

Seam Bias. Seam bias is the tendency of reported event transitions to occur in the first month of the reference months covered in an interview. For example, an individual who is asked in April to report her insurance status in each month since January may report being uninsured during all 4 months, even though her uninsured spell actually began in February. If that respondent had reported being insured in her interview the previous December, then her uninsured spell would be recorded as beginning in January—the "seam" month between the two interviews.

Seam bias is a well-known issue in the SIPP (Young 1989; Kalton, Hill, and Miller 1990; Czajka and Olsen 2000), which prompted the implementation of an extensive experimental research and development program at the Census Bureau during the 2001 SIPP panel (Moore et al. 2008). As a result of these efforts, the 2004 SIPP panel adopted dependent interviewing techniques in several domains of the survey, including the section on health insurance coverage.

Dependent interviewing is a surveying method to reduce seam bias that uses an individual's response from a previous interview as a starting point for later interviews. That is, in the example above, rather than simply ask about any insurance coverage in January through April, the surveyor would instead remind the respondent about her reported coverage in December, then ask if this coverage continued into January, February, and so on. In that way, rather than dividing time into separate blocks of reference months, the survey instrument maintains a more natural temporal sequence. The shift to dependent interviewing in the SIPP raises the question of how this change may affect estimates of uninsured spells; this is a question we take up in our results below.

By comparison, the MEPS has consistently used dependent interviewing for health insurance questions since the 1996 panel (Cohen 1997). Moreover, seam effects may also be less apparent in the MEPS because unlike in the SIPP—which spaces out interviews precisely every 4 months—interviews in the MEPS are more irregularly spaced. In the first year of MEPS Panel 10, for example, 13 percent of second interviews occurred after 4 months, 22 percent after 5 months, 30 percent after 6 months, and 13 percent after

7 months; alternative patterns are observed for other MEPS panels. Thus, to the extent seam bias issues affect MEPS estimates, their observed impact may be attenuated to some degree because the seam effects are more distributed across survey months.

Statistical Analysis

To investigate how the issues highlighted above manifest in both the SIPP and MEPS, we utilize nonparametric Kaplan–Meier survival methods to estimate the distribution of uninsured spell durations. One attractive feature of the Kaplan–Meier method is that it does not require parametric assumptions about the underlying distribution of hazards (Cox and Oakes 1984). This allows us to estimate quantities of research interest with minimal additional assumptions on the data-generating process.

Unless noted otherwise, all estimates are weighted using the calendar time weight corresponding to the month in which the individual's spell was reported to have begun. We also examine only incident uninsured spells and only one such spell per person. Standard errors and associated p-values are constructed using replicate weights to account for the complex sampling design of each survey.

RESULTS

Baseline Characteristics

Baseline characteristics of individuals with incident uninsured spells are shown in Table 1. These characteristics were chosen to reflect stable (e.g., race, gender) and easily predictable (e.g., age) attributes less prone to measurement differences between the two surveys.

Compared to adults in the MEPS, uninsured SIPP respondents were less white and more Hispanic. Interestingly, these differences were specific to incident uninsured respondents and were not observed when comparing *all* nonelderly adults and children across the two surveys (see Supplemental Appendix). There were also statistically significant differences in the age distribution among uninsured adults in the two surveys, though no systematic difference in age maintained over the three time periods. Among children there were also few sustained differences in baseline characteristics, though for 2001–2003 and 2004–2007 uninsured spells in the MEPS tended to be observed among older children.

Table 1: Baseline Characteristics of Uninsured Spells in the SIPP and MEPS, 2001–2003, 2004–2007, and 2008–2011

	2001–2003			2003–2007			2008–2011		
Nonelderly Adults	*SIPP n = 7,814 (%)*	*MEPS n = 3,395 (%)*	*p-value*	*SIPP n = 7,816 (%)*	*MEPS n = 4,513 (%)*	*p-value*	*SIPP n = 9,190 (%)*	*MEPS n = 3,736 (%)*	*p-value*
Age category (years)									
18–24	26.68	27.45	<.01*	27.56	28.02	<.01*	22.68	25.45	<.01*
25–34	27.37	26.16		25.36	24.6		26.71	26.42	
35–44	23.63	22.22		23.22	21.24		22.68	20.04	
45–54	15.96	15.24		17.85	17.34		20.58	16.75	
55–64	6.36	8.94		6.01	8.8		7.34	11.34	
Female	51.37	52.15	.44†	51.89	53.21	.25†	51.82	50.97	.46†
Married	36.88	37.36	.73†	34.92	37.36	.07†	41.75	35.72	<.01†
Race									
White, Non-Hispanic	60.90	63.25	.05*	56.54	65.00	<.01*	57.38	63.80	<.01*
Black, Non-Hispanic	15.10	14.98		15.59	13.47		15.17	14.67	
Other, Non-Hispanic	6.11	7.20		7.24	7.27		7.21	6.13	
Hispanic	17.89	14.56		20.63	14.26		20.24	15.39	
Census region									
Northeast	16.79	17.26	.43*	15.73	17.11	<.01*	16.17	15.40	.06*
Midwest	21.10	21.25		20.84	21.56		20.23	23.90	
South	36.56	36.28		36.08	36.58		37.37	35.91	
West	25.56	25.21		27.34	24.75		26.22	24.79	

Continued

Table 1. *Continued*

Children	*2001–2003* *SIPP* *n* = 4,513 (%)	*MEPS* *n* = 1,633 (%)	*p-value*	*2001–2003* *SIPP* *n* = 5,318 (%)	*MEPS* *n* = 2,214 (%)	*p-value*	*2001–2003* *SIPP* *n* = 6,311 (%)	*MEPS* *n* = 1,471 (%)	*p-value*
Age category (years)									
Under 12	72.17	65.17	<.01*	72.14	62.45	<.01*	71.35	61.56	<.01*
12–17	27.83	34.84		27.86	37.55		28.65	38.44	
Female	48.74	50.69	.24†	48.72	48.87	.93†	49.21	50.76	.36†
Race									
White, Non-Hispanic	54.00	52.93	.07*	51.16	51.56	.74*	44.59	50.82	.04*
Black, Non-Hispanic	19.36	15.43		14.63	13.40		16.56	13.26	
Other, Non-Hispanic	5.68	7.17		7.75	8.84		8.47	6.31	
Hispanic	20.96	24.47		26.45	26.21		30.38	29.60	
Census region									
Northeast	16.93	12.66	.17*	15.63	13.37	.19*	15.13	12.42	.45*
Midwest	19.68	18.75		19.95	17.67		20.05	21.94	
South	37.81	40.24		37.12	41.92		38.75	37.44	
West	25.58	28.35		27.31	27.04		26.07	28.21	

Notes. The above table reports baseline characteristics of freshly observed uninsured adults and children in the Medical Expenditure Panel Survey (MEPS) and the Survey of Income and Program Participation (SIPP).
*Chi-square goodness of fit.
†Test of proportion.
Sources: 2001, 2004, and 2008 SIPP and Panels 6–15 of the MEPS.

Uninsured Spell Durations

Figure 2 plots Kaplan–Meier curves of uninsured spell durations separately for each survey, study period, and population cohort. A striking feature is the sharp drop in the SIPP curves at each 4-month interval. Given previous evidence of seam bias in the SIPP, these drops likely reflect the tendency of previously uninsured respondents to report new coverage changes at the beginning of each 4-month reference period. Consequently, as individuals gain coverage, there is a tendency for spells to end after 4-month intervals.

Figure 2 also shows that seam transitions in the 2004–2007 SIPP curves are less pronounced than in 2001–2003, particularly among children. As discussed above, the 2004 SIPP incorporated dependent interviewing into the health insurance coverage questions. The visual evidence of shortened drops in the 2004 SIPP is consistent with U.S. Census Bureau findings that the adoption of dependent interviewing was associated with reductions in seam bias (Moore et al. 2008). However, Figure 2 also shows pronounced drops in the 2008 panel, indicating evidence of persistent seam effects even after these surveying changes.

Figure 2 also shows how nonparametric estimates of uninsured spell durations compare across the two surveys. As seen in Table 2, in the 2001–2003 and 2004–2007 periods, there are notable and statistically significant differences in the Kaplan–Meier curves ($p < .01$ for adults and $p < .01$ for chil-

Figure 2: Kaplan–Meier Curves by Survey, Year, and Population Cohort

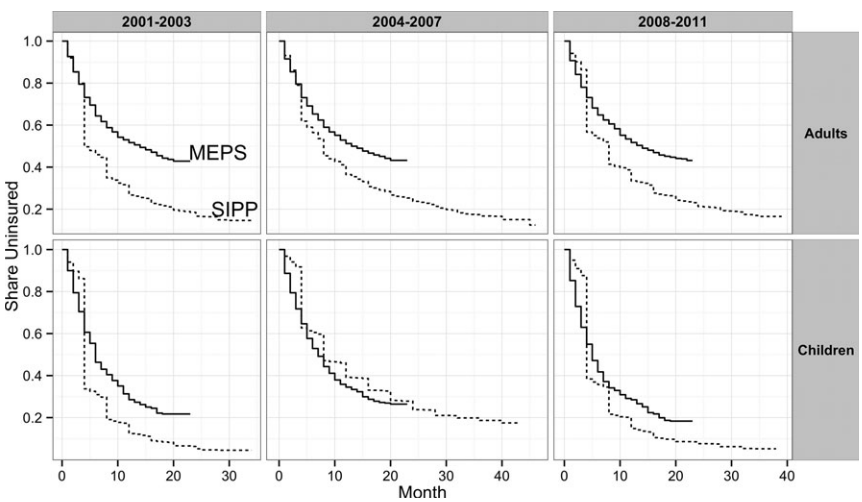


Table 2: Nonparametric Estimates of Uninsured Spell Durations in the SIPP and MEPS, 2001–2003, 2004–2007, and 2008–2011

	Cross-Sectional Weight					*Longitudinal Weight*				
	Quantile					*Quantile*				
	25%	*50%*	*75%*	*p-value Comparison Year: 2004*	*p-value MEPS versus SEPS*	*25%*	*50%*	*75%*	*p-value Comparison Year: 2004*	*p-value MEPS versus SEPS*
Nonelderly adults										
MEPS										
2001–2003	4	14	–*	.622	–‡	4	14	–*	.622	–‡
2004–2007	4	13	–*	–†	–‡	4	14	–*	–†	–‡
2008–2011	4	14	–*	.133	–‡	4	13	–*	.133	–‡
SIPP										
2001–2003	4	4	16	<.01	<.01	4	4	13	<.01	<.01
2004–2007	4	8	24	–†	<.01	4	8	21	–†	<.01
2008–2011	4	8	20	.041	<.01	4	7	17	.034	<.01
Children										
MEPS										
2001–2003	3	6	16	.356	–‡	3	6	16	.356	–‡
2004–2007	3	7	–*	–†	–‡	3	7	–*	–†	–‡
2008–2011	2	5	15	<.01	–‡	2	5	15	<.01	–‡
SIPP										
2001–2003	4	4	8	<.01	<.01	4	4	8	<.01	<.01
2004–2007	4	8	24	–†	<.01	4	8	24	–†	<.01
2008–2011	4	4	8	<.01	.633	4	4	8	<.01	.043

Notes. The above table reports nonparametric estimates of quantiles in the survival function for uninsured adults and children in the Medical Expenditure Panel Survey (MEPS) and the Survey of Income and Program Participation (SIPP).
*Nonparametric not available for this quantile.
†Comparison group for test of difference in survival function between 2001–2003, 2004–2007, 2008–2011 study periods.
‡Comparison group for test of difference in survival function between MEPS and SIPP within the study period.
Sources. 2001, 2004, and 2008 SIPP and Panels 6–15 of the MEPS.

dren), with shorter estimated spell durations in the SIPP. In 2008–2011, the survival curves for children are more similar (p = .633); however, among the adult sample in each time period a test of equality of curves is rejected at the alpha = .05 level.

Quantile Estimates

The set of results under the "Cross-Sectional Weight" heading in Table 2 summarize quantile estimates for the 25th, 50th, and 75th percentile of uninsured spell durations; as the heading implies, these estimates weight observations using the cross-sectional weight associated with the month in which their uninsured spell began. We estimate spells of about 4 months for adults and 2–4 months for children at the 25th percentile in both surveys.

The tendency of SIPP spell exits to occur in 4-month intervals results in shorter estimated median spell durations compared to the MEPS. For example, among adults in 2008–2011, we estimate median spell durations of 8 months using the SIPP, compared to a median of 14 months using the MEPS. As roughly 40 percent of uninsured adults in the MEPS had not gained coverage by the time their 24-month follow-up period ended, nonparametric estimates of the 75th percentile were not available for the adult MEPS cohorts. One exception is among children, where we estimate spell durations of 16 months at the 75th percentile using the MEPS for 2001–2003 and 15 months for 2008–2011, compared to 8 months over the same time periods using the SIPP.

Table 2 also shows how quantile estimates change over time. Notably, among MEPS adults we find little evidence of changes in the distribution of uninsured spells between 2001 and 2011. Among SIPP children, however, we find statistically significant changes, with longer spells observed in the 2004–2007 period.

Sensitivity to Weighting Method

The second panel of Table 2 reports estimated quantiles based on longitudinal weights. For the SIPP, most of the sensitivity to the weighting method occurs among adults in the upper quantiles of spell durations. At the 75th percentile for nonelderly adult spells, longitudinal weighting results in estimates that are 3 months shorter compared to the cross-sectional weight estimates.

While the 25th percentile and 50th percentile estimated durations are fairly robust to the weighting method used, an important consideration is the

Table 3: Sample Sizes by Survey and Weight

	SIPP		*MEPS*	
	Calendar	*Longitudinal*	*Calendar*	*Longitudinal*
Nonelderly adults				
2001–2003	7,814	6,708	3,395	3,395
2004–2007	7,816	6,333	4,513	4,513
2008–2011	9,190	6,834	3,736	3,736
Children				
2001–2003	4,513	4,002	1,633	1,633
2004–2007	5,318	4,488	2,214	2,214
2008–2011	6,311	4,638	1,471	1,471

loss in efficiency from using longitudinal weights, particularly in the SIPP. Table 3 summarizes the raw sample sizes under each weighting method. For example, due to survey attrition and administrative censoring in the 2004 (because of budget cuts), the longitudinally weighted sample for adults is only one-third the size of the sample using cross-sectional weights. This difference in effective sample sizes reduces the power of studies that elect to either restrict to fully observed individuals or that use longitudinal weights.

DISCUSSION

Understanding coverage dynamics is an important research goal, and the use of longitudinal household surveys will be important for understanding how these dynamics are changing in the United States. As we demonstrate here, the use of such surveys presents a unique set of methodological challenges and trade-offs. On the one hand, we find shorter estimated uninsured spell durations in the SIPP. The shorter SIPP durations may reflect seam bias, which results in estimated spell exits that may be recorded to occur earlier than they actually do. By comparison, we find that the influence of seam bias is less readily apparent in the MEPS. This likely reflects the fact that (a) MEPS has consistently adopted dependent interviewing methods to reduce seam bias; (b) has more irregularly spaced interviews; and (c) is sampled from among people with experience completing a comprehensive household survey.

While these factors may argue in favor of using the MEPS over the SIPP, a significant limitation of the MEPS is that nonparametric estimates were not possible due to the shorter longitudinal follow-up period. To the extent that public policy and research interest is more focused on better understanding long spells and designing policies to reduce them, then the SIPP may provide a better platform for understanding these dynamics, as such analyses are not possible with the MEPS without making strong parametric assumptions. Thus, a natural extension of our study would be to compare how parametric estimates differ across the two surveys.

Finally, we find that the choice of cross-sectional versus longitudinal weights does not materially affect estimates of spell durations for individuals in shorter spells, but it does so for individuals in long spells. Specifically, the use of longitudinal weights results in estimates that are 3 months shorter at the 75th percentile. As longitudinal weights are only defined for individuals in scope over the entire survey, this finding is consistent with the notion that these individuals may be in more stable situations and have systematically different traits that explain why they may experience shorter periods without health insurance. Again, to the extent research interest is more focused on reducing the incidence of uninsured spells among people in these situations, then analyses should consider the use of cross-sectional over longitudinal weights.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: We thank Katherine Swartz, Alan Zaslavsky, and Jonathan Gruber for their advice and insight on an earlier version of this manuscript. We gratefully acknowledge general research support from Vanderbilt University School of Medicine.

Disclosures: None.

Disclaimers: None.

NOTES

1. Though not due to nonresponse, the 2004 SIPP panel suffered an additional 58 percent loss in sample for the final four rounds due to budget cuts initiated by the U.S. Congress (as a compromise to the Bush administration's request that the SIPP be discontinued in 2007).
2. Differences in survey nonresponse could reflect differences in survey administration (over the phone for SIPP waves ≥2; in person for MEPS) and the fact that the MEPS respondents already have shown their willingness to participate in federal surveys (as each panel is sampled from the previous year's NHIS).
3. Moreover, accounting for these biases may require the use of parametric event history regression models, which in general require much stronger assumptions on the

distribution of the error term than standard linear regression (Cox and Oakes 1984; Cleves et al. 2008).

4. The distinction between left censoring and left truncation can be confusing because the terms are often used inconsistently in the econometrics and statistical literatures. For this paper, we adopt the (statistical literature) definition that truncation corresponds to the initiation (i.e., $t = 0$) of a given spell not being observed, while censoring (right and left) relates to the conclusion of the spell not being observed. That is, left truncation corresponds to observation of a spell being switched “on,” while right censoring corresponds to observation of a spell being switched “off” (Beyersmann, Allignol, and Schumacher 2012). Similarly, left censoring relates to spells that have switched “off” prior to observation.

References

Asgharian, M., C. E. M’Lan, and D. B. Wolfson. 2002. “Length-Biased Sampling with Right Censoring.” *Journal of American Statistical Association* 97: 201–9.

Beyersmann, J., A. Allignol, and M. Schumacher. 2012. *Competing Risks and Multistate Models with R.* New York: Springer.

Boudreau, C. 2003. *Duration Data Analysis in Longitudinal Surveys.* Waterloo, ON: University of Waterloo.

Call, K. T., M. Davern, and L. A. Blewett. 2007. “Estimates of Health Insurance Coverage: Comparing State Surveys with the Current Population Survey.” *Health Affairs (Millwood)* 26: 269–78.

Call, K. T., G. Davidson, M. Davern, and R. Nyman. 2008. “Medicaid Undercount and Bias to Estimates of Uninsurance: New Estimates and Existing Evidence.” *Health Services Research* 43: 901–14.

Carman, K., and C. Eibner. 2014. *Changes in Health Insurance Enrollment Since 2013: Evidence from the RAND Health Reform Opinion Study.* Santa Monica, CA: RAND Corporation.

Cleves, M., W. W. Gould, R. G. Gutierrez, and Y. Marchenko. 2008. *An Introduction to Survival Analysis Using Stata.* College Station, TX: Stata Press.

Cohen, J. 1997. *Methodology Report #1: Design and Methods of the Medical Expenditure Panel Survey Household Component.* Rockville, MD: Agency for Health Care Policy and Research.

Cox, D. R., and D. Oakes. 1984. *Analysis of Survival Data.* Boca Raton, FL: Chapman & Hall.

Cutler, D., and A. M. Gelber. 2009. “Changes in the Incidence and Duration of Periods without Insurance.” *New England Journal of Medicine* 360: 1740–8.

Czajka, J. L., J. Mabli, and S. Cody. 2008. *Sample Loss and Survey Bias in Estimates of Social Security Beneficiaries: A Table of Two Surveys.* Washington, DC: Mathematica Policy Research.

Czajka, J. L., and C. Olsen. 2000. *The Effects of Trigger Events on Changes in Children’s Health Insurance Coverage.* Washington, DC: Office of the Assistant Secretary for Planning and Evaluation (ASPE), Department of Health and Human Services.

Davern, M., and L. A. Blewett. 2007. "Distributing State Children's Health Insurance Program Funds: A Critical Review of the Design and Implementation of the Funding Formula." *Journal of Health Politics, Policy and Law* 32: 415–55.

Davern, M., J. Lepkowski, K. T. Call, N. Arnold, T. L. Johnson, K. Goldsteen, A. Todd-Malmlov, and L. A. Blewett. 2004. "Telephone Service Interruption Weighting Adjustments for State Health Insurance Surveys." *Inquiry* 41: 280–90.

Davern, M., H. Rodin, L. A. Blewett, and K. T. Call. 2007. "Are the Current Population Survey Uninsurance Estimates Too High? An Examination of the Imputation Process." *Health Services Research* 42: 2038–55.

Fronstin, P. 2010. *Sources of Health Insurance and Characteristics of the Uninsured: Analysis of the March 2010 Current Population Survey Sources of Health Insurance and Characteristics of the Uninsured: Analysis of the March 2010 Current Population Survey.*

Graves, J. A. 2012. "Medicaid Expansion Opt-Outs and Uncompensated Care." *New England Journal of Medicine* 367: 2365–7. doi:10.1056/NEJMp1209450.

Graves, J. A., and K. Swartz. 2012. "Health Care Reform and the Dynamics of Insurance Coverage—Lessons from Massachusetts." *New England Journal of Medicine* 367: 1181–4. doi:10.1056/NEJMp1207217.

———. 2013. "Understanding State Variation in Health Insurance Dynamics Can Help Tailor Enrollment Strategies for ACA Expansion." *Health Affairs (Millwood)* 32: 1832–40. doi:10.1377/hlthaff.2013.0327.

Kalton, G., D. Hill, and M. Miller. 1990. *The Seam Effect in Panel Surveys.* (Working Paper No. 116), Survey of Income and Progam Participation Working Paper Series. US Census Bureau and The Survey Research Center, University of Michigan.

Kauff, J., R. Olsen, and T. Fraker. 2002. *Nonrespondans and Nonresponse Bias: Evidence from a Survey of Former Welfare Recipients in Iowa* (No. 8217-909). Washington, DC: Mathematica Policy Research.

Klein, J. P., and M. L. Moeschberger. 2003. *Survival Analysis: Techniques for Censored and Truncated Data*, 2nd Edition. New York, NY: Springer.

Lancaster, T. 1992. *The Econometric Analysis of Transition Data.* Cambridge, England: Econometric Society.

Lohr, S. 1999. *Sampling: Design and Analysis.* Pacific Grove, CA: Duxbury Press.

Long, S. K., S. Zuckerman, and G. M. Kenney. 2014. *QuickTake: Number of Uninsured Adults Falls By 5.4 Million Since 2013.* Washington, DC: The Urban Institute.

Moore, J., N. Bates, J. Pascale, and A. Okon. 2008. *Tackling Seam Bias through Questionnaire Design: Methodology of Longitudinal Surveys.* Chichester, England: John Wiley and Sons.

Short, P. F., D. Graefe, and C. Schoen. 2003. *Churn, Churn, Churn: How Instability of Health Insurance Shapes America's Uninsured Problem.* New York, NY: The Commonwealth Fund.

Sommers, B. D., T. Musco, K. Finegold, M. Z. Gunja, A. Burke, and A. M. McDowell. 2014. "Health Reform and Changes in Health Insurance Coverage in 2014." *New England Journal of Medicine* 371: 867–74. doi:10.1056/NEJMsr1406753.

State Health Access Data Assistance Center. 2010. *Comparing Federal Government Surveys that Count the Uninsured.* Minneapolis, MN: Robert Wood Johnson Foundation.

Swartz, K. 1986. "Interpreting the Estimates from Four National Surveys of the Number of People without Health Insurance." *Journal of Economic and Social Measurement* 14: 233–42.

Swartz, K., and T. McBride. 1990. "Spells without Health Insurance: Distributions of Durations and Their Link to Point-in-Time Estimates of the Uninsured." *Inquiry* 27: 281–8.

Swartz, K., T. McBride, and J. Marcotte. 1993. "Spells without Health Insurance: The Distribution of Durations When Left-Censored Spells Are Included." *Inquiry* 30: 77–83.

de Una-Alvarez, J. 2002. "Product-Limit Estimation for Length-Biased Censored Data." *Test* 11: 109–25.

———. 2004. "Nonparametric Estimation under Length-Biased Sampling and Type I Censoring: A Moment Based Approach." *Annals of the Institute of Statistical Mathematics* 56: 667.

U.S. Census Bureau. 2001. *Survey of Income and Program Participation.* Washington, DC: U.S. Census Bureau.

U.S. Census Bureau. 2008a. *SIPP User's Guide.* Washington, DC: U.S. Census Bureau.

U.S. Census Bureau. 2008b. *2008 Survey of Income and Program Particpation Technical Documentation.* Washington, DC: U.S. Census Bureau.

U.S. Census Bureau. 2009. *Reengineering the Survey of Income and Program Participation.* Washington, DC: National Academies Press.

Wolfson, C., D. B. Wolfson, M. Asgharian, C. E. M'Lan, T. Østbye, K. Rockwood, and D. B. Hogan. 2001. "A Reevaluation of the Duration of Survival after the Onset of Dementia." *New England Journal of Medicine* 344: 1111–6. doi:10.1056/NEJM200104123441501.

Young, N. 1989. *Wave Seam Effects (No. 98), Survey of Income and Progam Participation Working Paper Series.* Washington, DC: The Urban Institute.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Table A1: Baseline Characteristics of Uninsured Spells in the SIPP and MEPS, 2001–2003, 2004–2007, and 2008–2011.

Table A2: Nonparametric Estimates of Uninsured Spell Durations in the First Year of SIPP and MEPS, 2001–2003 and 2004–2007.